

# BAB I

## PENDAHULUAN

### A. Latar Belakang Masalah

Dalam kehidupan, seringkali peneliti dihadapkan dengan suatu kejadian yang saling berhubungan atau berpengaruh satu sama lain. Ilmu statistika mengenal metode regresi yang merupakan salah satu metode untuk menentukan tingkat pengaruh suatu variabel terhadap variabel yang lain. Variabel yang diteliti adalah variabel independen dan variabel dependen. Variabel independen yaitu variabel bebas yang tidak dipengaruhi oleh variabel lainnya dan dinotasikan dengan variabel  $X$ . Variabel dependen yaitu variabel yang keberadaannya dipengaruhi oleh variabel bebas dan dinotasikan dengan variabel  $Y$ .

Regresi pertama kali dikemukakan oleh Francis Galton, seorang antropolog dan ahli meteorologi Perancis dalam artikelnya yang berjudul “*Family Likeness in Stature*”. Tetapi ada pula yang menyatakan istilah regresi muncul pada pidato Francis Galton di depan Section H of The British Association di Aberdeen tahun 1855 yang dimuat dalam makalah *Regression Toward Mediocrity in Hereditary Stature* (Draper and Smith, 1985). Galton menunjukkan bahwa tinggi badan anak laki-laki dari ayah setelah beberapa generasi cenderung mundur (*regressed*) mendekati nilai tengah populasi. Istilah regresi sekarang diterapkan pada semua jenis peramalan, dan tidak harus berimplikasi suatu regresi mendekati nilai tengah populasi.

Untuk memodelkan hubungan antara variabel dependen dan independen maka dilakukan analisis regresi. Analisis regresi adalah suatu analisis yang digunakan untuk mengukur pengaruh antara variabel independen dan variabel dependen (Danang Sunyoto, 2010). Arnita (2013) mengatakan bahwa terdapat beberapa tujuan penggunaan analisis regresi, yaitu: (a) Membuat estimasi rata-rata dan nilai variabel dependen dengan didasari pada nilai variabel independen, (b) Menguji hipotesis karakteristik dependensi, (c) Untuk meramalkan nilai rata-rata variabel independen dengan didasarkan pada nilai variabel bebas di luar jangkauan sampel.

Terdapat dua jenis regresi yaitu regresi linear dan regresi non linear. Regresi linear yaitu regresi yang menyatakan bentuk hubungan dimana variabel dependen dan independennya berpangkat satu. Regresi non linear adalah regresi yang menyatakan bentuk hubungan dimana variabel dependen dan independennya mempunyai pangkat tertentu.

Berdasarkan jumlah variabel independennya ( $X$ ), analisis regresi linear dapat dibedakan menjadi dua macam, yaitu analisis regresi linear sederhana dan analisis regresi linear ganda. Analisis regresi linear sederhana adalah analisis regresi untuk mengetahui hubungan antara satu variabel independen ( $X$ ) dengan variabel dependen ( $Y$ ) saja. Sedangkan analisis regresi linear ganda adalah analisis regresi untuk mengetahui hubungan antara dua atau lebih variabel independen ( $X$ ) dengan variabel dependen ( $Y$ ). Analisis regresi linear ganda lebih sering digunakan karena suatu keadaan biasanya dipengaruhi oleh beberapa keadaan lainnya.

Teknik estimasi variabel dependen yang mendasari analisis regresi klasik adalah metode kuadrat terkecil/*Least Square Method*. Metode ini pertama kali diperkenalkan oleh Carl Friederich Gauss. Untuk mendapatkan nilai-nilai penduga parameter, metode kuadrat terkecil memerlukan beberapa asumsi klasik yang harus dipenuhi yaitu memenuhi normalitas, tidak terjadi autokorelasi, tidak terjadi heteroskedastisitas, dan tidak terjadi multikolinearitas. Apabila keempat asumsi tersebut terpenuhi, maka penduga parameter yang diperoleh bersifat *Best Linier Unbiased Estimator (BLUE)*. Namun tidak jarang pula ditemui hal-hal yang menyebabkan asumsi-asumsi tersebut tidak terpenuhi, sehingga penggunaan metode kuadrat terkecil akan memberikan kesimpulan yang kurang baik atau nilai penduga parameternya bersifat bias.

Salah satu pelanggaran asumsi klasik adalah terjadinya multikolinearitas. Multikolinearitas adalah kondisi dimana terdapat hubungan linier atau korelasi yang tinggi antara masing-masing variabel independen dalam model regresi ganda. Model regresi yang baik seharusnya tidak terjadi korelasi di antara variabel bebas. Adanya korelasi yang tinggi diantara variabel bebas akan menghasilkan penaksir yang berbias, memiliki varians yang besar, dan mungkin jauh dari nilai sasaran. Oleh karena itu, multikolinearitas harus diatasi karena dianggap sebagai suatu kelemahan dalam analisis regresi.

Dalam upaya mengatasi masalah multikolinearitas, muncullah beberapa teknik regresi modern yang dapat digunakan. Salah satu cara untuk mengatasi masalah tersebut adalah melalui metode *Ridge Regression*. *Ridge Regression*

pertama kali dikemukakan oleh Hoerl and Kennard pada tahun 1970. Pada dasarnya metode ini juga merupakan metode kuadrat terkecil. Perbedaannya adalah bahwa pada metode *Ridge Regression*, nilai variabel bebasnya ditransformasikan dahulu melalui prosedur *centering and rescaling*.

Pada tahun 1970, Hoerl and Kennard membahas tentang estimasi *Ridge Regression* dalam penelitiannya yang berjudul *Ridge Regression: Biased Estimation for Non orthogonal Problems*. Penelitian ini menjelaskan tentang estimator *Ordinary Ridge Regression (ORR)* dan bagaimana menambah tetapan bias  $k$  dengan menggunakan metode *ridge trace* untuk mendapatkan perkiraan bias dengan *Mean Squared Error (MSE)* lebih kecil. Pada tahun 2011, M. El-Dereny dan N.I. Rashwan dalam jurnalnya *Solving Multicollinearity Problem Using Ridge Regression Models* menjelaskan tentang beberapa metode untuk mendapatkan estimasi pada analisis regresi *ridge*, diantaranya *Generalized Ridge Regression (GRR)*, *Ordinary Ridge Regression (ORR)*, dan *Directed Ridge Regression (DRR)*. Jurnal ini menunjukkan bahwa penduga atau estimator dari regresi *ridge* lebih baik dibanding estimator *Ordinary Least Square (OLS)* saat ditemukannya kasus multikolinearitas.

Cara lain untuk mengatasi multikolinearitas dikemukakan oleh Toutenbuerg (1982), yaitu dengan menggunakan informasi prior pada ruang parameter. Informasi prior yang dimaksud berbentuk persamaan restriksi linear, yaitu:

$$\mathbf{R}\boldsymbol{\beta} = \mathbf{r} \text{ atau } \boldsymbol{\delta} = \mathbf{R}\boldsymbol{\beta} - \mathbf{r} \quad (1.1)$$

dengan:

$\mathbf{R}$  menyatakan matriks berukuran  $n \times h$  yang menunjukkan struktur informasi pada parameter  $\boldsymbol{\beta}$ , baik secara individual atau kombinasi linear dari elemen-elemen vektor  $\boldsymbol{\beta}$ .

$\boldsymbol{\beta}$  menyatakan vektor parameter berukuran  $h \times 1$

$\mathbf{r}$  menyatakan vektor restriksi berukuran  $n \times 1$ .

$\boldsymbol{\delta} = \mathbf{0}$  apabila restriksi linear benar ( $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ ) dan  $\boldsymbol{\delta} \neq \mathbf{0}$  apabila restriksi linear salah ( $\mathbf{R}\boldsymbol{\beta} \neq \mathbf{r}$ )

Selanjutnya, Gollberger dalam Toutenburg (1982) mengkombinasikan restriksi linear dengan *OLS* sehingga diperoleh estimator baru yang disebut *Restricted Least Square (RLS)*. Sifat dari *RLS* tergantung pada restriksi linear, artinya apabila restriksi linear benar atau  $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$  maka estimator yang diperoleh bersifat tak bias. Apabila restriksi linear salah atau  $\mathbf{R}\boldsymbol{\beta} \neq \mathbf{r}$  estimator yang diperoleh bersifat bias. Kelebihan dari penggunaan restriksi linear ini yaitu dapat menurunkan variansi estimator. Secara keseluruhan, estimator *RLS* mempunyai nilai *Mean Square Error (MSE)* lebih kecil dari estimator *OLS*. Akan tetapi, penentuan restriksi linear tersebut belum pasti benar atau salah sehingga dalam penggunaannya kadang kala timbul ketidakpastian. Apabila kondisi restriksi linear benar atau  $\mathbf{R}\boldsymbol{\beta} = \mathbf{r}$  maka kadang kala parameter masih memuat persoalan multikolinearitas. Oleh karena itu, akan dibahas suatu estimator baru yang memodifikasi estimator *RLS* dengan arah estimator *Ridge Regression*.

Nityananda Sarkar A. memperkenalkan *Restricted Ridge Regression (RRR)* dalam jurnal *A new estimator combining the ridge regression and the restricted least squares methods of estimation*, pada tahun 1992. Dalam jurnal ini dijelaskan bahwa *RRR* merupakan kombinasi antara *Ordinary Ridge Regression (ORR)* dan *Restricted Least Square (RLS)*. Jurnal ini juga menunjukkan bahwa *RRR* adalah estimator bias dan mempunyai nilai *MSE* lebih kecil dari *ORR* dan *RLS*.

Pada tahun 2012, S. Najarian, M. Arashi, dan B.M. Golam Kibria dalam jurnalnya yang berjudul *a simulation study on some restricted ridge regression estimator* membahas tentang simulasi penggunaan estimator *RRR* pada beberapa contoh data, seperti *Acetylene Data*, *Portland Cement Data*, dan *Gross National Product Data*, serta membandingkan nilai *MSE* dari *RRR* dan *RLS*. Hasil dari simulasi adalah estimator *RRR* lebih baik dari *RLS* dengan melihat nilai *MSE RRR* yang lebih kecil dari *MSE RLS*.

Beberapa penelitian dengan tema serupa juga telah dilakukan. Soemartini (2008) dalam tesisnya yang berjudul “*penyelesaian multikolinearitas melalui metode ridge regression*” menggunakan metode *ridge regression* untuk mengatasi terjadinya multikolinearitas dalam regresi linear ganda. Metode yang dibahas pada dasarnya juga merupakan metode kuadrat terkecil, perbedaannya adalah bahwa pada metode *ridge regression*, nilai variabel *regressornya* ditransformasikan dahulu melalui prosedur *centering and rescaling*. Estira Woro Astrini (2013) dalam skripsinya “*analisis regresi ridge dua tahap untuk permasalahan multikolinearitas*”. Analisis

regresi *Ridge* dua tahap ini merupakan gabungan antara *Two Stage Least Square (TSLS)* dengan *Ordinary Ridge Regression (ORR)*.

Alzaber (1998) juga menjelaskan estimator koefisien regresi dalam tesisnya yang berjudul “*kombinasi estimator regresi ridge dan estimator kudrat terkecil restriksi*”. Dalam tesisnya ditunjukkan keunggulan *estimator regresi ridge restriksi* bila dibandingkan dengan *estimator kudrat terkecil restriksi dan estimator regresi ridge* dengan menggunakan kriteria rata-rata sesatan kuadrat.

Masalah multikolinearitas terutama terjadi pada variabel-variabel yang memuat data *time series*, dimana korelasi antar variabel independennya tinggi. Untuk memperjelas penggunaan estimator *RRR* dalam mengatasi multikolinearitas, akan dibahas salah satu contoh kasus dimana pada variabel bebasnya memuat multikolinearitas. Salah satu contohnya adalah pada uang primer. Uang primer adalah uang yang dikeluarkan oleh Bank Indonesia yang kemudian uang tersebut didistribusikan kepada Bank Umum, Bank Perkreditan Rakyat (BPR), dan sektor swasta. Terdapat beberapa faktor yang mempengaruhi besarnya uang primer, diantaranya adalah tagihan kepada bukan penduduk, kredit likuiditas, serta tagihan kepada bank umum dan BPR.

Pada data uang primer, uang primer sebagai variabel dependen (*Y*) sedangkan faktor-faktor yang mempengaruhi uang primer sebagai variabel independen (*X*). Dalam kasus ini, diduga terjadi pelanggaran asumsi multikolinearitas diantara variabel-variabel independennya, yaitu variabel kredit likuiditas dan tagihan kepada bank umum dan BPR. Uang primer

merupakan variabel moneter yang cukup penting dan digunakan sebagai indikator bagi kebijakan moneter terhadap perekonomian, maka penyelesaian kasus multikolinearitas dilakukan tanpa menghilangkan salah satu variabel independennya. Oleh karena itu, digunakan metode *Restricted Ridge Regression* (*RRR*) untuk penyelesaian kasus multikolinearitas yang terjadi.

Dalam tugas akhir ini, akan dibahas langkah-langkah untuk estimasi *Restricted Ridge Regression* (*RRR*) yang merupakan kombinasi dari *Restricted Least Square* (*RLS*) dan *Ordinary Ridge Regression* (*ORR*) dan penerapan *RRR* untuk mengatasi masalah multikolinearitas yang terjadi pada data uang primer.

## **B. Batasan Masalah**

Agar penyelesaian masalah tidak menyimpang dari pembahasan, maka perlu dibuat pembatasan masalah yaitu dianggap bahwa asumsi klasik yang lain tetap terpenuhi sehingga difokuskan pada permasalahan multikolinearitas dan cara penanganan pelanggaran asumsi tersebut dengan *Restricted Ridge Regression* (*RRR*).

## **C. Perumusan Masalah**

Berdasarkan latar belakang di atas maka yang menjadi pokok permasalahan adalah :

1. Bagaimana langkah-langkah estimasi *Restricted Ridge Regression* (*RRR*)?
2. Bagaimana penerapan *Restricted Ridge Regression* (*RRR*) dalam mengatasi masalah multikolinearitas?



#### **D. Tujuan Penelitian**

Berdasarkan permasalahan di atas, tujuan dari penulisan tugas akhir ini adalah :

1. Mengetahui langkah-langkah estimasi *Restricted Ridge Regression (RRR)*.
2. Mengetahui penerapan *Restricted Ridge Regression (RRR)* dalam mengatasi masalah multikolinearitas.

#### **E. Manfaat Penelitian**

Manfaat penulisan tugas akhir ini diantaranya:

1. Bagi Penulis

Membuka wawasan teoritis maupun praktis pada kajian ilmu *Ridge Regression*, khususnya *Restricted Ridge Regression* dalam mengatasi masalah multikolinearitas.

2. Bagi Jurusan Matematika FMIPA UNY

Menambah kelengkapan koleksi pustaka dan menjadi referensi untuk penelitian-penelitian selanjutnya.